

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: SHYJAN, Andrew
- (ii) TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED
POLYPEPTIDE
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Testa, Hurwitz & Thibault
(B) STREET: 125 High St.
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02110
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: WALLER, Patrick R H
(B) REGISTRATION NUMBER: 41,418
(C) REFERENCE/DOCKET NUMBER: MIL-001CP
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 248-7000
(B) TELEFAX: (617) 248-7100

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4847 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 116..4426

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGCTCATGCT CGGGAGCGTG GTTGAGCGGC TGGCGCGGTT GTCCTGGAGC AGGGGCGCAG 60
GAATTCTGAT GTGAAGCTAA CAGTCTGTGA GCCCTGGAAC CTCCACTCAG AGAAG ATG 118
Met
1
AAG GAT ATC GAC ATA GGA AAA GAG TAT ATC ATC CCC AGT CCT GGG TAT 166
Lys Asp Ile Asp Ile Gly Lys Glu Tyr Ile Ile Pro Ser Pro Gly Tyr
5 10 15
AGA AGT GTG AGG GAG AGA ACC AGC ACT TCT GGG ACG CAC AGA GAC CGT 214
Arg Ser Val Arg Glu Arg Thr Ser Thr Ser Gly Thr His Arg Asp Arg
20 25 30
GAA GAT TCC AAG TTC AGG AGA ACT CGA CCG TTG GAA TGC CAA GAT GCC 262
Glu Asp Ser Lys Phe Arg Arg Thr Arg Pro Leu Glu Cys Gln Asp Ala
35 40 45
TTG GAA ACA GCA GCC CGA GCC GAG GGC CTC TCT CTT GAT GCC TCC ATG 310
Leu Glu Thr Ala Ala Arg Ala Glu Gly Leu Ser Leu Asp Ala Ser Met
50 55 60 65
CAT TCT CAG CTC AGA ATC CTG GAT GAG GAG CAT CCC AAG GGA AAG TAC 358
His Ser Gln Leu Arg Ile Leu Asp Glu Glu His Pro Lys Gly Lys Tyr
70 75 80
CAT CAT GGC TTG AGT GCT CTG AAG CCC ATC CGG ACT ACT TGC AAA CAC 406
His His Gly Leu Ser Ala Leu Lys Pro Ile Arg Thr Thr Cys Lys His
85 90 95
CAG CAC CCA GTG GAC AAT GCT GGG CTT TTT TCC TGT ATG ACT TTT TCG 454
Gln His Pro Val Asp Asn Ala Gly Leu Phe Ser Cys Met Thr Phe Ser
100 105 110
TGG CTT TCT TCT CTG GCC CGT GTG GCC CAC AAG AAG GGG GAG CTC TCA 502
Trp Leu Ser Ser Leu Ala Arg Val Ala His Lys Lys Gly Glu Leu Ser
115 120 125
ATG GAA GAC GTG TGG TCT CTG TCC AAG CAC GAG TCT TCT GAC GTG AAC 550
Met Glu Asp Val Trp Ser Leu Ser Lys His Glu Ser Ser Asp Val Asn
130 135 140 145
TGC AGA AGA CTA GAG AGA CTG TGG CAA GAA GAG CTG AAT GAA GTT GGG 598
Cys Arg Arg Leu Glu Arg Leu Trp Gln Glu Glu Leu Asn Glu Val Gly
150 155 160
CCA GAC GCT GCT TCC CTG CGA AGG GTT GTG TGG ATC TTC TGC CGC ACC 646
Pro Asp Ala Ala Ser Leu Arg Arg Val Val Trp Ile Phe Cys Arg Thr
165 170 175
AGG CTC ATC CTG TCC ATC GTG TGC CTG ATG ATC ACG CAG CTG GCT GGC 694
Arg Leu Ile Leu Ser Ile Val Cys Leu Met Ile Thr Gln Leu Ala Gly
180 185 190
TTC AGT GGA CCA GCC TTC ATG GTG AAA CAC CTC TTG GAG TAT ACC CAG 742
Phe Ser Gly Pro Ala Phe Met Val Lys His Leu Leu Glu Tyr Thr Gln
195 200 205

GCA ACA GAG TCT AAC CTG CAG TAC AGC TTG TTG TTA GTG CTG GGC CTC Ala Thr Glu Ser Asn Leu Gln Tyr Ser Leu Leu Leu Val Leu Gly Leu 210 215 220 225	790
CTC CTG ACG GAA ATC GTG CGG TCT TGG TCG CTT GCA CTG ACT TGG GCA Leu Leu Thr Glu Ile Val Arg Ser Trp Ser Leu Ala Leu Thr Trp Ala 230 235 240	838
TTG AAT TAC CGA ACC GGT GTC CGC TTG CGG GGG GCC ATC CTA ACC ATG Leu Asn Tyr Arg Thr Gly Val Arg Leu Arg Gly Ala Ile Leu Thr Met 245 250 255	886
GCA TTT AAG AAG ATC CTT AAG TTA AAG AAC ATT AAA GAG AAA TCC CTG Ala Phe Lys Lys Ile Leu Lys Leu Lys Asn Ile Lys Glu Lys Ser Leu 260 265 270	934
GGT GAG CTC ATC AAC ATT TGC TCC AAC GAT GGG CAG AGA ATG TTT GAG Gly Glu Leu Ile Asn Ile Cys Ser Asn Asp Gly Gln Arg Met Phe Glu 275 280 285	982
GCA GCA GCC GTT GGC AGC CTG CTG GCT GGA GGA CCC GTT GTT GCC ATC Ala Ala Ala Val Gly Ser Leu Leu Ala Gly Gly Pro Val Val Ala Ile 290 295 300 305	1030
TTA GGC ATG ATT TAT AAT GTA ATT ATT CTG GGA CCA ACA GGC TTC CTG Leu Gly Met Ile Tyr Asn Val Ile Ile Leu Gly Pro Thr Gly Phe Leu 310 315 320	1078
GGA TCA GCT GTT TTT ATC CTC TTT TAC CCA GCA ATG ATG TTT GCA TCA Gly Ser Ala Val Phe Ile Leu Phe Tyr Pro Ala Met Met Phe Ala Ser 325 330 335	1126
CGG CTC ACA GCA TAT TTC AGG AGA AAA TGC GTG GCC GCC ACG GAT GAA Arg Leu Thr Ala Tyr Phe Arg Arg Lys Cys Val Ala Ala Thr Asp Glu 340 345 350	1174
CGT GTC CAG AAG ATG AAT GAA GTT CTT ACT TAC ATT AAA TTT ATC AAA Arg Val Gln Lys Met Asn Glu Val Leu Thr Tyr Ile Lys Phe Ile Lys 355 360 365	1222
ATG TAT GCC TGG GTC AAA GCA TTT TCT CAG AGT GTT CAG AAA ATC CGC Met Tyr Ala Trp Val Lys Ala Phe Ser Gln Ser Val Gln Lys Ile Arg 370 375 380 385	1270
GAG GAG GAG CGT CGG ATA TTG GAA AAA GCC GGG TAC TTC CAG AGC ATC Glu Glu Glu Arg Arg Ile Leu Glu Lys Ala Gly Tyr Phe Gln Ser Ile 390 395 400	1318
ACT GTG GGT GTG GCT CCC ATT GTG GTG GTG ATT GCC AGC GTG GTG ACC Thr Val Gly Val Ala Pro Ile Val Val Val Ile Ala Ser Val Val Thr 405 410 415	1366
TTC TCT GTT CAT ATG ACC CTG GGC TTC GAT CTG ACA GCA GCA CAG GCT Phe Ser Val His Met Thr Leu Gly Phe Asp Leu Thr Ala Ala Gln Ala 420 425 430	1414
TTC ACA GTG GTG ACA GTC TTC AAT TCC ATG ACT TTT GCT TTG AAA GTA Phe Thr Val Val Thr Val Phe Asn Ser Met Thr Phe Ala Leu Lys Val 435 440 445	1462

ACA CCG TTT TCA GTA AAG TCC CTC TCA GAA GCC TCA GTG GCT GTT GAC Thr Pro Phe Ser Val Lys Ser Leu Ser Glu Ala Ser Val Ala Val Asp 450 455 460 465	1510
AGA TTT AAG AGT TTG TTT CTA ATG GAA GAG GTT CAC ATG ATA AAG AAC Arg Phe Lys Ser Leu Phe Leu Met Glu Glu Val His Met Ile Lys Asn 470 475 480	1558
AAA CCA GCC AGT CCT CAC ATC AAG ATA GAG ATG AAA AAT GCC ACC TTG Lys Pro Ala Ser Pro His Ile Lys Ile Glu Met Lys Asn Ala Thr Leu 485 490 495	1606
GCA TGG GAC TCC TCC CAC TCC AGT ATC CAG AAC TCG CCC AAG CTG ACC Ala Trp Asp Ser Ser His Ser Ser Ile Gln Asn Ser Pro Lys Leu Thr 500 505 510	1654
CCC AAA ATG AAA AAA GAC AAG AGG GCT TCC AGG GGC AAG AAA GAG AAG Pro Lys Met Lys Lys Asp Lys Arg Ala Ser Arg Gly Lys Lys Glu Lys 515 520 525	1702
GTG AGG CAG CTG CAG CGC ACT GAG CAT CAG GCG GTG CTG GCA GAG CAG Val Arg Gln Leu Gln Arg Thr Glu His Gln Ala Val Leu Ala Glu Gln 530 535 540 545	1750
AAA GGC CAC CTC CTC CTG GAC AGT GAC GAG CGG CCC AGT CCC GAA GAG Lys Gly His Leu Leu Leu Asp Ser Asp Glu Arg Pro Ser Pro Glu Glu 550 555 560	1798
GAA GAA GGC AAG CAC ATC CAC CTG GGC CAC CTG CGC TTA CAG AGG ACA Glu Glu Gly Lys His Ile His Leu Gly His Leu Arg Leu Gln Arg Thr 565 570 575	1846
CTG CAC AGC ATC GAT CTG GAG ATC CAA GAG GGT AAA CTG GTT GGA ATC Leu His Ser Ile Asp Leu Glu Ile Gln Glu Gly Lys Leu Val Gly Ile 580 585 590	1894
TGC GGC AGT GTG GGA AGT GGA AAA ACC TCT CTC ATT TCA GCC ATT TTA Cys Gly Ser Val Gly Ser Gly Lys Thr Ser Leu Ile Ser Ala Ile Leu 595 600 605	1942
GGC CAG ATG ACG CTT CTA GAG GGC AGC ATT GCA ATC AGT GGA ACC TTC Gly Gln Met Thr Leu Leu Glu Gly Ser Ile Ala Ile Ser Gly Thr Phe 610 615 620 625	1990
GCT TAT GTG GCC CAG CAG GCC TGG ATC CTC AAT GCT ACT CTG AGA GAC Ala Tyr Val Ala Gln Gln Ala Trp Ile Leu Asn Ala Thr Leu Arg Asp 630 635 640	2038
AAC ATC CTG TTT GGG AAG GAA TAT GAT GAA GAA AGA TAC AAC TCT GTG Asn Ile Leu Phe Gly Lys Glu Tyr Asp Glu Glu Arg Tyr Asn Ser Val 645 650 655	2086
CTG AAC AGC TGC TGC CTG AGG CCT GAC CTG GCC ATT CTT CCC AGC AGC Leu Asn Ser Cys Cys Leu Arg Pro Asp Leu Ala Ile Leu Pro Ser Ser 660 665 670	2134
GAC CTG ACG GAG ATT GGA GAG CGA GGA GCC AAC CTG AGC GGT GGG CAG Asp Leu Thr Glu Ile Gly Glu Arg Gly Ala Asn Leu Ser Gly Gly Gln 675 680 685	2182

CGC CAG AGG ATC AGC CTT GCC CGG GCC TTG TAT AGT GAC AGG AGC ATC Arg Gln Arg Ile Ser Leu Ala Arg Ala Leu Tyr Ser Asp Arg Ser Ile 690 695 700 705	2230
TAC ATC CTG GAC GAC CCC CTC AGT GCC TTA GAT GCC CAT GTG GGC AAC Tyr Ile Leu Asp Asp Pro Leu Ser Ala Leu Asp Ala His Val Gly Asn 710 715 720	2278
CAC ATC TTC AAT AGT GCT ATC CGG AAA CAT CTC AAG TCC AAG ACA GTT His Ile Phe Asn Ser Ala Ile Arg Lys His Leu Lys Ser Lys Thr Val 725 730 735	2326
CTG TTT GTT ACC CAC CAG TTA CAG TAC CTG GTT GAC TGT GAT GAA GTG Leu Phe Val Thr His Gln Leu Gln Tyr Leu Val Asp Cys Asp Glu Val 740 745 750	2374
ATC TTC ATG AAA GAG GGC TGT ATT ACG GAA AGA GGC ACC CAT GAG GAA Ile Phe Met Lys Glu Gly Cys Ile Thr Glu Arg Gly Thr His Glu Glu 755 760 765	2422
CTG ATG AAT TTA AAT GGT GAC TAT GCT ACC ATT TTT AAT AAC CTG TTG Leu Met Asn Leu Asn Gly Asp Tyr Ala Thr Ile Phe Asn Asn Leu Leu 770 775 780 785	2470
CTG GGA GAG ACA CCG CCA GTT GAG ATC AAT TCA AAA AAG GAA ACC AGT Leu Gly Glu Thr Pro Pro Val Glu Ile Asn Ser Lys Lys Glu Thr Ser 790 795 800	2518
GGT TCA CAG AAG AAG TCA CAA GAC AAG GGT CCT AAA ACA GGA TCA ATA Gly Ser Gln Lys Lys Ser Gln Asp Lys Gly Pro Lys Thr Gly Ser Ile 805 810 815	2566
AAG AAG GAA AAA GCA GTA AAG CCA GAG GAA GGG CAG CTT GTG CAG CTG Lys Lys Glu Lys Ala Val Lys Pro Glu Glu Gly Gln Leu Val Gln Leu 820 825 830	2614
GAA GAG AAA GGG CAG GGT TCA GTG CCC TGG TCA GTA TAT GGT GTC TAC Glu Glu Lys Gly Gln Gly Ser Val Pro Trp Ser Val Tyr Gly Val Tyr 835 840 845	2662
ATC CAG GCT GCT GGG GGC CCC TTG GCA TTC CTG GTT ATT ATG GCC CTT Ile Gln Ala Ala Gly Gly Pro Leu Ala Phe Leu Val Ile Met Ala Leu 850 855 860 865	2710
TTC ATG CTG AAT GTA GGC AGC ACC GCC TTC AGC ACC TGG TGG TTG AGT Phe Met Leu Asn Val Gly Ser Thr Ala Phe Ser Thr Trp Trp Leu Ser 870 875 880	2758
TAC TGG ATC AAG CAA GGA AGC GGG AAC ACC ACT GTG ACT CGA GGG AAC Tyr Trp Ile Lys Gln Gly Ser Gly Asn Thr Thr Val Thr Arg Gly Asn 885 890 895	2806
GAG ACC TCG GTG AGT GAC AGC ATG AAG GAC AAT CCT CAT ATG CAG TAC Glu Thr Ser Val Ser Asp Ser Met Lys Asp Asn Pro His Met Gln Tyr 900 905 910	2854
TAT GCC AGC ATC TAC GCC CTC TCC ATG GCA GTC ATG CTG ATC CTG AAA Tyr Ala Ser Ile Tyr Ala Leu Ser Met Ala Val Met Leu Ile Leu Lys 915 920 925	2902

GCC ATT CGA GGA GTT GTC TTT GTC AAG GGC ACG CTG CGA GCT TCC TCC Ala Ile Arg Gly Val Val Phe Val Lys Gly Thr Leu Arg Ala Ser Ser 930 935 940 945	2950
CGG CTG CAT GAC GAG CTT TTC CGA AGG ATC CTT CGA AGC CCT ATG AAG Arg Leu His Asp Glu Leu Phe Arg Arg Ile Leu Arg Ser Pro Met Lys 950 955 960	2998
TTT TTT GAC ACG ACC CCC ACA GGG AGG ATT CTC AAC AGG TTT TCC AAA Phe Phe Asp Thr Thr Pro Thr Gly Arg Ile Leu Asn Arg Phe Ser Lys 965 970 975	3046
GAC ATG GAT GAA GTT GAC GTG CGG CTG CCG TTC CAG GCC GAG ATG TTC Asp Met Asp Glu Val Asp Val Arg Leu Pro Phe Gln Ala Glu Met Phe 980 985 990	3094
ATC CAG AAC GTT ATC CTG GTG TTC TTC TGT GTG GGA ATG ATC GCA GGA Ile Gln Asn Val Ile Leu Val Phe Phe Cys Val Gly Met Ile Ala Gly 995 1000 1005	3142
GTC TTC CCG TGG TTC CTT GTG GCA GTG GGG CCC CTT GTC ATC CTC TTT Val Phe Pro Trp Phe Leu Val Ala Val Gly Pro Leu Val Ile Leu Phe 1010 1015 1020 1025	3190
TCA GTC CTG CAC ATT GTC TCC ACG GTC CTG ATT CGG GAG CTG AAG CGT Ser Val Leu His Ile Val Ser Arg Val Leu Ile Arg Glu Leu Lys Arg 1030 1035 1040	3238
CTG GAC AAT ATC ACG CAG TCA CCT TTC CTC TCC CAC ATC ACG TCC AGC Leu Asp Asn Ile Thr Gln Ser Pro Phe Leu Ser His Ile Thr Ser Ser 1045 1050 1055	3286
ATA CAG GGC CTT GCC ACC ATC CAC GCC TAC AAT AAA GGG CAG GAG TTT Ile Gln Gly Leu Ala Thr Ile His Ala Tyr Asn Lys Gly Gln Glu Phe 1060 1065 1070	3334
CTG CAC AGA TAC CAG GAG CTG CTG GAT GAC AAC CAA GCT CCT TTT TTT Leu His Arg Tyr Gln Glu Leu Leu Asp Asp Asn Gln Ala Pro Phe Phe 1075 1080 1085	3382
TTG TTT ACG TGT GCG ATG CGG TGG CTG GCT GTG CGG CTG GAC CTC ATC Leu Phe Thr Cys Ala Met Arg Trp Leu Ala Val Arg Leu Asp Leu Ile 1090 1095 1100 1105	3430
AGC ATC GCC CTC ATC ACC ACC ACG GGG CTG ATG ATC GTT CTT ATG CAC Ser Ile Ala Leu Ile Thr Thr Thr Gly Leu Met Ile Val Leu Met His 1110 1115 1120	3478
GGG CAG ATT CCC CCA GCC TAT GCG GGT CTC GCC ATC TCT TAT GCT GTC Gly Gln Ile Pro Pro Ala Tyr Ala Gly Leu Ala Ile Ser Tyr Ala Val 1125 1130 1135	3526
CAG TTA ACG GGG CTG TTC CAG TTT ACG GTC AGA CTG GCA TCT GAG ACA Gln Leu Thr Gly Leu Phe Gln Phe Thr Val Arg Leu Ala Ser Glu Thr 1140 1145 1150	3574
GAA GCT CGA TTC ACC TCG GTG GAG AGG ATC AAT CAC TAC ATT AAG ACT Glu Ala Arg Phe Thr Ser Val Glu Arg Ile Asn His Tyr Ile Lys Thr 1155 1160 1165	3622

CTG TCC TTG GAA GCA CCT GCC AGA ATT AAG AAC AAG GCT CCC TCC CCT Leu Ser Leu Glu Ala Pro Ala Arg Ile Lys Asn Lys Ala Pro Ser Pro 1170 1175 1180 1185	3670
GAC TGG CCC CAG GAG GGA GAG GTG ACC TTT GAG AAC GCA GAG ATG AGG Asp Trp Pro Gln Glu Gly Glu Val Thr Phe Glu Asn Ala Glu Met Arg 1190 1195 1200	3718
TAC CGA GAA AAC CTC CCT CTC GTC CTA AAG AAA GTA TCC TTC ACG ATC Tyr Arg Glu Asn Leu Pro Leu Val Leu Lys Lys Val Ser Phe Thr Ile 1205 1210 1215	3766
AAA CCT AAA GAG AAG ATT GGC ATT GTG GGG CGG ACA GGA TCA GGG AAG Lys Pro Lys Glu Lys Ile Gly Ile Val Gly Arg Thr Gly Ser Gly Lys 1220 1225 1230	3814
TCC TCG CTG GGG ATG GCC CTC TTC CGT CTG GTG GAG TTA TCT GGA GGC Ser Ser Leu Gly Met Ala Leu Phe Arg Leu Val Glu Leu Ser Gly Gly 1235 1240 1245	3862
TGC ATC AAG ATT GAT GGA GTG AGA ATC AGT GAT ATT GGC CTT GCC GAC Cys Ile Lys Ile Asp Gly Val Arg Ile Ser Asp Ile Gly Leu Ala Asp 1250 1255 1260 1265	3910
CTC CGA AGC AAA CTC TCT ATC ATT CCT CAA GAG CCG GTG CTG TTC AGT Leu Arg Ser Lys Leu Ser Ile Ile Pro Gln Glu Pro Val Leu Phe Ser 1270 1275 1280	3958
GGC ACT GTC AGA TCA AAT TTG GAC CCC TTC AAC CAG TAC ACT GAA GAC Gly Thr Val Arg Ser Asn Leu Asp Pro Phe Asn Gln Tyr Thr Glu Asp 1285 1290 1295	4006
CAG ATT TGG GAT GCC CTG GAG AGG ACA CAC ATG AAA GAA TGT ATT GCT Gln Ile Trp Asp Ala Leu Glu Arg Thr His Met Lys Glu Cys Ile Ala 1300 1305 1310	4054
CAG CTA CCT CTG AAA CTT GAA TCT GAA GTG ATG GAG AAT GGG GAT AAC Gln Leu Pro Leu Lys Leu Glu Ser Glu Val Met Glu Asn Gly Asp Asn 1315 1320 1325	4102
TTC TCA GTG GGG GAA CGG CAG CTC TTG TGC ATA GCT AGA GCC CTG CTC Phe Ser Val Gly Glu Arg Gln Leu Leu Cys Ile Ala Arg Ala Leu Leu 1330 1335 1340 1345	4150
CGC CAC TGT AAG ATT CTG ATT TTA GAT GAA GCC ACA GCT GCC ATG GAC Arg His Cys Lys Ile Leu Ile Leu Asp Glu Ala Thr Ala Ala Met Asp 1350 1355 1360	4198
ACA GAG ACA GAC TTA TTG ATT CAA GAG ACC ATC CGA GAA GCA TTT GCA Thr Glu Thr Asp Leu Leu Ile Gln Glu Thr Ile Arg Glu Ala Phe Ala 1365 1370 1375	4246
GAC TGT ACC ATG CTG ACC ATT GCC CAT CGC CTG CAC ACG GTT CTA GGC Asp Cys Thr Met Leu Thr Ile Ala His Arg Leu His Thr Val Leu Gly 1380 1385 1390	4294
TCC GAT AGG ATT ATG GTG CTG GCC CAG GGA CAG GTG GTG GAG TTT GAC Ser Asp Arg Ile Met Val Leu Ala Gln Gly Gln Val Val Glu Phe Asp 1395 1400 1405	4342

ACC CCA TCG GTC CTT CTG TCC AAC GAC AGT TCC CGA TTC TAT GCC ATG 4390
 Thr Pro Ser Val Leu Leu Ser Asn Asp Ser Ser Arg Phe Tyr Ala Met
 1410 1415 1420 1425

TTT GCT GCT GCA GAG AAC AAG GTC GCT GTC AAG GGC TGACTCCTCC 4436
 Phe Ala Ala Ala Glu Asn Lys Val Ala Val Lys Gly
 1430 1435

CTGTTGACGA AGTCTCTTTT CTTTAGAGCA TTGCCATTCC CTGCCTGGGG CGGGCCCCCTT 4496

CATCGCGTCC TCCTACCGAA ACCTTGCCCTT TCTCGATTTT ATCTTTCGCA CAGCAGTTCC 4556

GGATTGGCTT GTGTGTTTCA CTTTATAGGGA GAGTCATATT TTGATTATTG TATTTATTCC 4616

ATATTGATGT AAACAAAATT TAGTTTTTGT TCTTAATTGC ACTCTAAAAG GTTCAGGGAA 4676

CCGTTATTAT AATTGTATCA GAGGCCTATA ATGAAGCTTT ATACGTGTAG CTATATCTAT 4736

ATATAATTCT GTACATAGCC TATATTTTACA GTGAAAATGT AAGCTGTTTA TTTTATATTA 4796

AAATAAGCAC TGTGCTAAAA AAAAAAAAAA AAAAAAAAAA AGGGCGGCCG C 4847

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1437 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Lys Asp Ile Asp Ile Gly Lys Glu Tyr Ile Ile Pro Ser Pro Gly
 1 5 10 15

Tyr Arg Ser Val Arg Glu Arg Thr Ser Thr Ser Gly Thr His Arg Asp
 20 25 30

Arg Glu Asp Ser Lys Phe Arg Arg Thr Arg Pro Leu Glu Cys Gln Asp
 35 40 45

Ala Leu Glu Thr Ala Ala Arg Ala Glu Gly Leu Ser Leu Asp Ala Ser
 50 55 60

Met His Ser Gln Leu Arg Ile Leu Asp Glu Glu His Pro Lys Gly Lys
 65 70 75 80

Tyr His His Gly Leu Ser Ala Leu Lys Pro Ile Arg Thr Thr Cys Lys
 85 90 95

His Gln His Pro Val Asp Asn Ala Gly Leu Phe Ser Cys Met Thr Phe
 100 105 110

Ser Trp Leu Ser Ser Leu Ala Arg Val Ala His Lys Lys Gly Glu Leu
 115 120 125

Ser Met Glu Asp Val Trp Ser Leu Ser Lys His Glu Ser Ser Asp Val
 130 135 140

Asn Cys Arg Arg Leu Glu Arg Leu Trp Gln Glu Glu Leu Asn Glu Val
 145 150 155 160
 Gly Pro Asp Ala Ala Ser Leu Arg Arg Val Val Trp Ile Phe Cys Arg
 165 170 175
 Thr Arg Leu Ile Leu Ser Ile Val Cys Leu Met Ile Thr Gln Leu Ala
 180 185 190
 Gly Phe Ser Gly Pro Ala Phe Met Val Lys His Leu Leu Glu Tyr Thr
 195 200 205
 Gln Ala Thr Glu Ser Asn Leu Gln Tyr Ser Leu Leu Leu Val Leu Gly
 210 215 220
 Leu Leu Leu Thr Glu Ile Val Arg Ser Trp Ser Leu Ala Leu Thr Trp
 225 230 235 240
 Ala Leu Asn Tyr Arg Thr Gly Val Arg Leu Arg Gly Ala Ile Leu Thr
 245 250 255
 Met Ala Phe Lys Lys Ile Leu Lys Leu Lys Asn Ile Lys Glu Lys Ser
 260 265 270
 Leu Gly Glu Leu Ile Asn Ile Cys Ser Asn Asp Gly Gln Arg Met Phe
 275 280 285
 Glu Ala Ala Ala Val Gly Ser Leu Leu Ala Gly Gly Pro Val Val Ala
 290 295 300
 Ile Leu Gly Met Ile Tyr Asn Val Ile Ile Leu Gly Pro Thr Gly Phe
 305 310 315 320
 Leu Gly Ser Ala Val Phe Ile Leu Phe Tyr Pro Ala Met Met Phe Ala
 325 330 335
 Ser Arg Leu Thr Ala Tyr Phe Arg Arg Lys Cys Val Ala Ala Thr Asp
 340 345 350
 Glu Arg Val Gln Lys Met Asn Glu Val Leu Thr Tyr Ile Lys Phe Ile
 355 360 365
 Lys Met Tyr Ala Trp Val Lys Ala Phe Ser Gln Ser Val Gln Lys Ile
 370 375 380
 Arg Glu Glu Glu Arg Arg Ile Leu Glu Lys Ala Gly Tyr Phe Gln Ser
 385 390 395 400
 Ile Thr Val Gly Val Ala Pro Ile Val Val Val Ile Ala Ser Val Val
 405 410 415
 Thr Phe Ser Val His Met Thr Leu Gly Phe Asp Leu Thr Ala Ala Gln
 420 425 430
 Ala Phe Thr Val Val Thr Val Phe Asn Ser Met Thr Phe Ala Leu Lys
 435 440 445
 Val Thr Pro Phe Ser Val Lys Ser Leu Ser Glu Ala Ser Val Ala Val
 450 455 460

Asp Arg Phe Lys Ser Leu Phe Leu Met Glu Glu Val His Met Ile Lys
 465 470 475 480
 Asn Lys Pro Ala Ser Pro His Ile Lys Ile Glu Met Lys Asn Ala Thr
 485 490 495
 Leu Ala Trp Asp Ser Ser His Ser Ser Ile Gln Asn Ser Pro Lys Leu
 500 505 510
 Thr Pro Lys Met Lys Lys Asp Lys Arg Ala Ser Arg Gly Lys Lys Glu
 515 520 525
 Lys Val Arg Gln Leu Gln Arg Thr Glu His Gln Ala Val Leu Ala Glu
 530 535 540
 Gln Lys Gly His Leu Leu Leu Asp Ser Asp Glu Arg Pro Ser Pro Glu
 545 550 555 560
 Glu Glu Glu Gly Lys His Ile His Leu Gly His Leu Arg Leu Gln Arg
 565 570 575
 Thr Leu His Ser Ile Asp Leu Glu Ile Gln Glu Gly Lys Leu Val Gly
 580 585 590
 Ile Cys Gly Ser Val Gly Ser Gly Lys Thr Ser Leu Ile Ser Ala Ile
 595 600 605
 Leu Gly Gln Met Thr Leu Leu Glu Gly Ser Ile Ala Ile Ser Gly Thr
 610 615 620
 Phe Ala Tyr Val Ala Gln Gln Ala Trp Ile Leu Asn Ala Thr Leu Arg
 625 630 635 640
 Asp Asn Ile Leu Phe Gly Lys Glu Tyr Asp Glu Glu Arg Tyr Asn Ser
 645 650 655
 Val Leu Asn Ser Cys Cys Leu Arg Pro Asp Leu Ala Ile Leu Pro Ser
 660 665 670
 Ser Asp Leu Thr Glu Ile Gly Glu Arg Gly Ala Asn Leu Ser Gly Gly
 675 680 685
 Gln Arg Gln Arg Ile Ser Leu Ala Arg Ala Leu Tyr Ser Asp Arg Ser
 690 695 700
 Ile Tyr Ile Leu Asp Asp Pro Leu Ser Ala Leu Asp Ala His Val Gly
 705 710 715 720
 Asn His Ile Phe Asn Ser Ala Ile Arg Lys His Leu Lys Ser Lys Thr
 725 730 735
 Val Leu Phe Val Thr His Gln Leu Gln Tyr Leu Val Asp Cys Asp Glu
 740 745 750
 Val Ile Phe Met Lys Glu Gly Cys Ile Thr Glu Arg Gly Thr His Glu
 755 760 765
 Glu Leu Met Asn Leu Asn Gly Asp Tyr Ala Thr Ile Phe Asn Asn Leu
 770 775 780

Leu Leu Gly Glu Thr Pro Pro Val Glu Ile Asn Ser Lys Lys Glu Thr
 785 790 795 800
 Ser Gly Ser Gln Lys Lys Ser Gln Asp Lys Gly Pro Lys Thr Gly Ser
 805 810 815
 Ile Lys Lys Glu Lys Ala Val Lys Pro Glu Glu Gly Gln Leu Val Gln
 820 825 830
 Leu Glu Glu Lys Gly Gln Gly Ser Val Pro Trp Ser Val Tyr Gly Val
 835 840 845
 Tyr Ile Gln Ala Ala Gly Gly Pro Leu Ala Phe Leu Val Ile Met Ala
 850 855 860
 Leu Phe Met Leu Asn Val Gly Ser Thr Ala Phe Ser Thr Trp Trp Leu
 865 870 875 880
 Ser Tyr Trp Ile Lys Gln Gly Ser Gly Asn Thr Thr Val Thr Arg Gly
 885 890 895
 Asn Glu Thr Ser Val Ser Asp Ser Met Lys Asp Asn Pro His Met Gln
 900 905 910
 Tyr Tyr Ala Ser Ile Tyr Ala Leu Ser Met Ala Val Met Leu Ile Leu
 915 920 925
 Lys Ala Ile Arg Gly Val Val Phe Val Lys Gly Thr Leu Arg Ala Ser
 930 935 940
 Ser Arg Leu His Asp Glu Leu Phe Arg Arg Ile Leu Arg Ser Pro Met
 945 950 955 960
 Lys Phe Phe Asp Thr Thr Pro Thr Gly Arg Ile Leu Asn Arg Phe Ser
 965 970 975
 Lys Asp Met Asp Glu Val Asp Val Arg Leu Pro Phe Gln Ala Glu Met
 980 985 990
 Phe Ile Gln Asn Val Ile Leu Val Phe Phe Cys Val Gly Met Ile Ala
 995 1000 1005
 Gly Val Phe Pro Trp Phe Leu Val Ala Val Gly Pro Leu Val Ile Leu
 1010 1015 1020
 Phe Ser Val Leu His Ile Val Ser Arg Val Leu Ile Arg Glu Leu Lys
 1025 1030 1035 1040
 Arg Leu Asp Asn Ile Thr Gln Ser Pro Phe Leu Ser His Ile Thr Ser
 1045 1050 1055
 Ser Ile Gln Gly Leu Ala Thr Ile His Ala Tyr Asn Lys Gly Gln Glu
 1060 1065 1070
 Phe Leu His Arg Tyr Gln Glu Leu Leu Asp Asp Asn Gln Ala Pro Phe
 1075 1080 1085
 Phe Leu Phe Thr Cys Ala Met Arg Trp Leu Ala Val Arg Leu Asp Leu
 1090 1095 1100

Ile Ser Ile Ala Leu Ile Thr Thr Thr Gly Leu Met Ile Val Leu Met
 1105 1110 1115 1120
 His Gly Gln Ile Pro Pro Ala Tyr Ala Gly Leu Ala Ile Ser Tyr Ala
 1125 1130 1135
 Val Gln Leu Thr Gly Leu Phe Gln Phe Thr Val Arg Leu Ala Ser Glu
 1140 1145 1150
 Thr Glu Ala Arg Phe Thr Ser Val Glu Arg Ile Asn His Tyr Ile Lys
 1155 1160 1165
 Thr Leu Ser Leu Glu Ala Pro Ala Arg Ile Lys Asn Lys Ala Pro Ser
 1170 1175 1180
 Pro Asp Trp Pro Gln Glu Gly Glu Val Thr Phe Glu Asn Ala Glu Met
 1185 1190 1195 1200
 Arg Tyr Arg Glu Asn Leu Pro Leu Val Leu Lys Lys Val Ser Phe Thr
 1205 1210 1215
 Ile Lys Pro Lys Glu Lys Ile Gly Ile Val Gly Arg Thr Gly Ser Gly
 1220 1225 1230
 Lys Ser Ser Leu Gly Met Ala Leu Phe Arg Leu Val Glu Leu Ser Gly
 1235 1240 1245
 Gly Cys Ile Lys Ile Asp Gly Val Arg Ile Ser Asp Ile Gly Leu Ala
 1250 1255 1260
 Asp Leu Arg Ser Lys Leu Ser Ile Ile Pro Gln Glu Pro Val Leu Phe
 1265 1270 1275 1280
 Ser Gly Thr Val Arg Ser Asn Leu Asp Pro Phe Asn Gln Tyr Thr Glu
 1285 1290 1295
 Asp Gln Ile Trp Asp Ala Leu Glu Arg Thr His Met Lys Glu Cys Ile
 1300 1305 1310
 Ala Gln Leu Pro Leu Lys Leu Glu Ser Glu Val Met Glu Asn Gly Asp
 1315 1320 1325
 Asn Phe Ser Val Gly Glu Arg Gln Leu Leu Cys Ile Ala Arg Ala Leu
 1330 1335 1340
 Leu Arg His Cys Lys Ile Leu Ile Leu Asp Glu Ala Thr Ala Ala Met
 1345 1350 1355 1360
 Asp Thr Glu Thr Asp Leu Leu Ile Gln Glu Thr Ile Arg Glu Ala Phe
 1365 1370 1375
 Ala Asp Cys Thr Met Leu Thr Ile Ala His Arg Leu His Thr Val Leu
 1380 1385 1390
 Gly Ser Asp Arg Ile Met Val Leu Ala Gln Gly Gln Val Val Glu Phe
 1395 1400 1405
 Asp Thr Pro Ser Val Leu Leu Ser Asn Asp Ser Ser Arg Phe Tyr Ala
 1410 1415 1420

Met Phe Ala Ala Ala Glu Asn Lys Val Ala Val Lys Gly
1425 1430 1435

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 463 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCGTCCGCCT AGAACGCAGA GATGAGCTAC CGAGAAAACC TCCCTCTCGT CCTAAAGAAA 60
GTATCCTTCA CGATCAAACC TAAAGAGAAAG ATTGGCATTG TGGGGCGGAC AGGATCAGGG 120
AAGTCCTCGC TGGGGATGGC CCTCTTCCGT CTGGTGGAGT TATCTGGAGG CTGCATCAAG 180
ATTGATGGAG TGAGAATCAG TGATATTGGC CTTGCCGACC TCCGAAGCAA ACTCTCTATC 240
ATTCCTCAAG AGCCGGTGCT GTTCAGTGGC ACTGTCAGAT CAAATTTGGA CCCTTCAACC 300
AGTACACTGA AGACCAGATT TGGGATGCCC TGGAAAGGAC ACACATGAAA GAATGTATTG 360
CTCCAGCTAC CTCCTGAAAC TTGAATCCTG AATTTGATGG AGAAATGGGG AAATAACTTC 420
TCCAGTTGGG GGAAACGGCA CTCTTTGTTG CCATACCTAN ACC 463

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TGCTGGTTCT CTCCTCACA CTTC 24

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTCGGCTCGG GCTGCTGTTT CCAA

24

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGGTGCTGGT GTTTGGAAGT AGTC

24

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCCAGAGAAG AAAGCCACGA AAAA

24

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGGCACACGA TGGACAGGAT GAGC

24